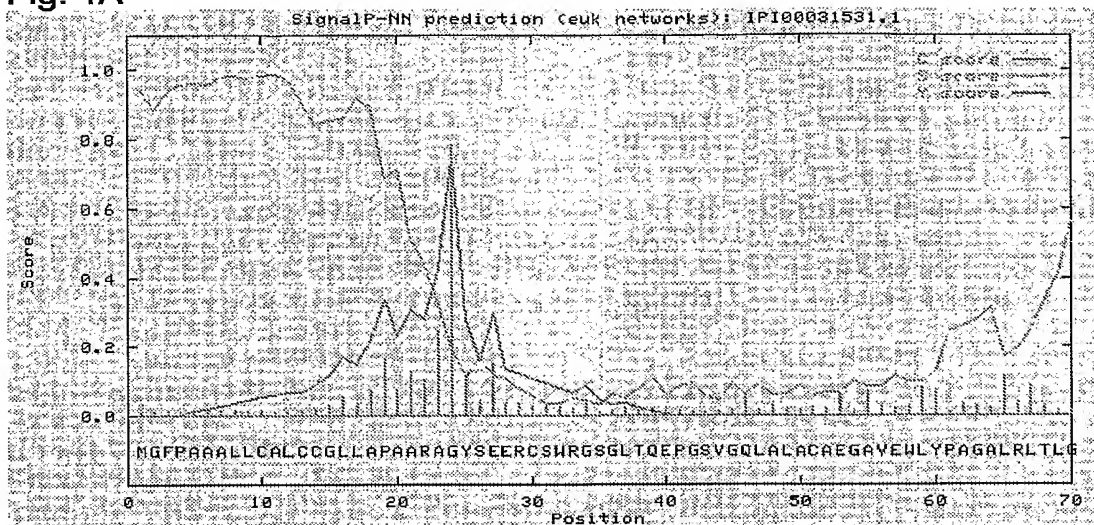


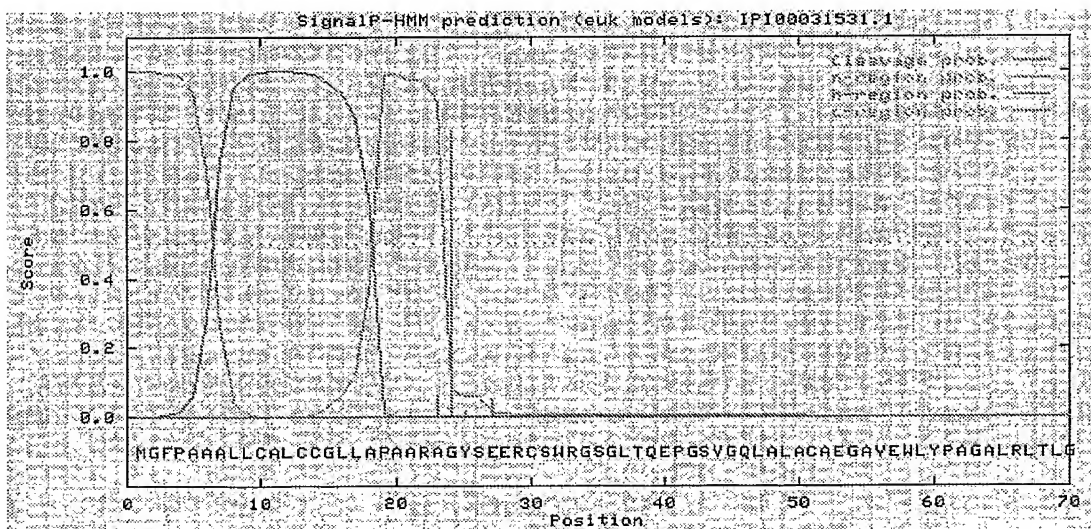
1/13

Fig. 1A



```
>IPI00031531.1          length = 293
# Measure  Position  Value  Cutoff  signal peptide?
  max. C      24      0.785  0.33   YES
  max. Y      24      0.746  0.32   YES
  max. S      11      0.989  0.82   YES
  mean S     1-23     0.846  0.47   YES
# Most likely cleavage site between pos. 23 and 24: ARA-GY
```

Fig. 1B



```
>IPI00031531.1
Prediction: Signal peptide
Signal peptide probability: 1.000
Signal anchor probability: 0.000
Max cleavage site probability: 0.832 between pos. 23 and 24
```

2/13

FIG. 2**human NsG33 (SEQ ID No 3)**

# Gene Ontology category	Odds
Signal_transducer	0.538
Receptor	0.433
Hormone	1.173
Structural_protein	0.168
Transporter	0.230
Ion_channel	0.372
Voltage-gated_ion_channel	0.130
Cation_channel	0.215
Transcription	0.294
Transcription_regulation	0.152
Stress_response	0.340
Immune_response	0.186
Growth_factor	2.083
Metal_ion_transport	0.020

human N-terminal peptide (SEQ ID No 19)

# Gene Ontology category	Odds
Signal_transducer	0.464
Receptor	0.296
Hormone	0.206
Structural_protein	0.987
Transporter	0.311
Ion_channel	0.147
Voltage-gated_ion_channel	0.157
Cation_channel	0.215
Transcription	0.311
Transcription_regulation	0.829
Stress_response	0.162
Immune_response	1.460
Growth_factor	8.142
Metal_ion_transport	0.061

human C-terminal peptide (SEQ ID No 5)

# Gene Ontology category	Odds
Signal_transducer	0.242
Receptor	0.038
Hormone	0.303
Structural_protein	0.096
Transporter	0.231
Ion_channel	0.185
Voltage-gated_ion_channel	0.191
Cation_channel	0.215
Transcription	0.312
Transcription_regulation	0.295
Stress_response	0.145
Immune_response	0.157
Growth_factor	7.963
Metal_ion_transport	0.020

3/13

Fig 3a. CLUSTAL W (1.82) multiple sequence alignment

```

Mouse NsG33      HASAHASALLCALCCGLLAASAHAGYSEDRCSWRGSGLTQEPGSGVGLTLTDCTEGAIEWL
Rat NsG33       ---MLVAALLCALCCGLLAASARAGYSEDRCSWRGSGLTQEPGSGVGLTLTDCTEGAIEWL
Human NsG33     -MGFPAAALLCALCCGLLAPARAGYSEERCSSWRGSGLTQEPGSGVGLLALACAEGAVEWL
                  .:*****.:*:*****:*****:*****:* *:***:***

Mouse NsG33      YPAGALRLTLGGPDPGTRPSIVCLRPFRPFAGAQVFAERMTGNLELLLAEGPDLAGGRCM
Rat NsG33       YPAGALRLTLGGSDPGTRPSIVCLRPTRPFAGAQVFAERMAGNLELLLAEGQGLAGGRCM
Human NsG33     YPAGALRLTLGGPDPRARPGIACLRPVRPFAGAQVFAERAGGALELLLAEGPGPAGGRCV
                  *****.*:*.*.**** ***** * ***** . *****:

Mouse NsG33      RWGPRRRALFLQATPHRDISRRVAAFREFELHEDQRAEMSPQAQGLGVDGACRPCSDAEL
Rat NsG33       RWGPRRRALFLQATPHRDISRRVAAFQFELHEDQRAEMSPQAQGFVDGACRPCSDAEL
Human NsG33     RWGPRRRALFLQATPHQDISRRVAAFREFELREDGRPELPPQAHGLGVDGACRPCSDAEL
                  *****:*****:***:* *.*:.**:*:*****

Mouse NsG33      LLAactsdfVIHGtiHGVahdTELQESVITVVVARVIRQTLPLFKESGSEGGQGRASIRTL
Rat NsG33       LLTACTSDFVIHGtiHGVVHDMELQESVITVVATRIVIRQTLPLFQESGSEGRGQASVRTL
Human NsG33     LLAactsdfVIHGtiHGVTHDVELQESVITVVAARVLRQTPPLFQAGRSGDQGLTSIRTP
                  *:***** *****.* *****.:*:*** ***: * * .:* :*:**

Mouse NsG33      LRCGVRPGPGSFLFMGWSRFGEAWLGCAPRFQEFSSRVYSAALTTHLNPCEMALD
Rat NsG33       LRCGVRPGPGSFLFMGWSRFGEAWLGCAPRFQEFSSRVYSAALAAHLNPCEVALD
Human NsG33     LRCGVHPGPGTFLFMGWSRFGEARLGCAPRFQEFRRAYEAAARAHLHPCEVALH
                  *****.**:***** ***** *.*.*:*.***:*.

```

Fig 3b. CLUSTAL W (1.82) multiple sequence alignment

```

Mouse      --MLVATLLCALCCGLLAASAHAGYSEDRCSWRGSGLTQEPGSGVGLTLTDCTEGAIEWLY 58
Rat        --MLVAALLCALCCGLLAASARAGYSEDRCSWRGSGLTQEPGSGVGLTLTDCTEGAIEWLY 58
Human      MGFPAAALLCALCCGLLAPARAGYSEERCSSWRGSGLTQEPGSGVGLLALACAEGAVEWLY 60
            .:*****.:*:*****:*****:*****:* *:***:***

Mouse      PAGALRLTLGGPDPGTRPSIVCLRPFRPFAGAQVFAERMTGNLELLLAEGPDLAGGRCMR 118
Rat        PAGALRLTLGGSDPGTRPSIVCLRPTRPFAGAQVFAERMAGNLELLLAEGQGLAGGRCMR 118
Human      PAGALRLTLGGPDPRARPGIACLRPVRPFAGAQVFAERAGGALELLLAEGPGPAGGRCVR 120
            *****.*:*.*.**** ***** * ***** . *****:

Mouse      WGPERRRALFLQATPHRDISRRVAAFREFELHEDQRAEMSPQAQGLGVDGACRPCSDAELL 178
Rat        WGPERRRALFLQATPHRDISRRVAAFQFELHEDQRAEMSPQAQGFVDGACRPCSDAELL 178
Human      WGPERRRALFLQATPHQDISRRVAAFREFELREDGRPELPPQAHGLGVDGACRPCSDAELL 180
            *****:*****:***:* *.*:.**:*:*****

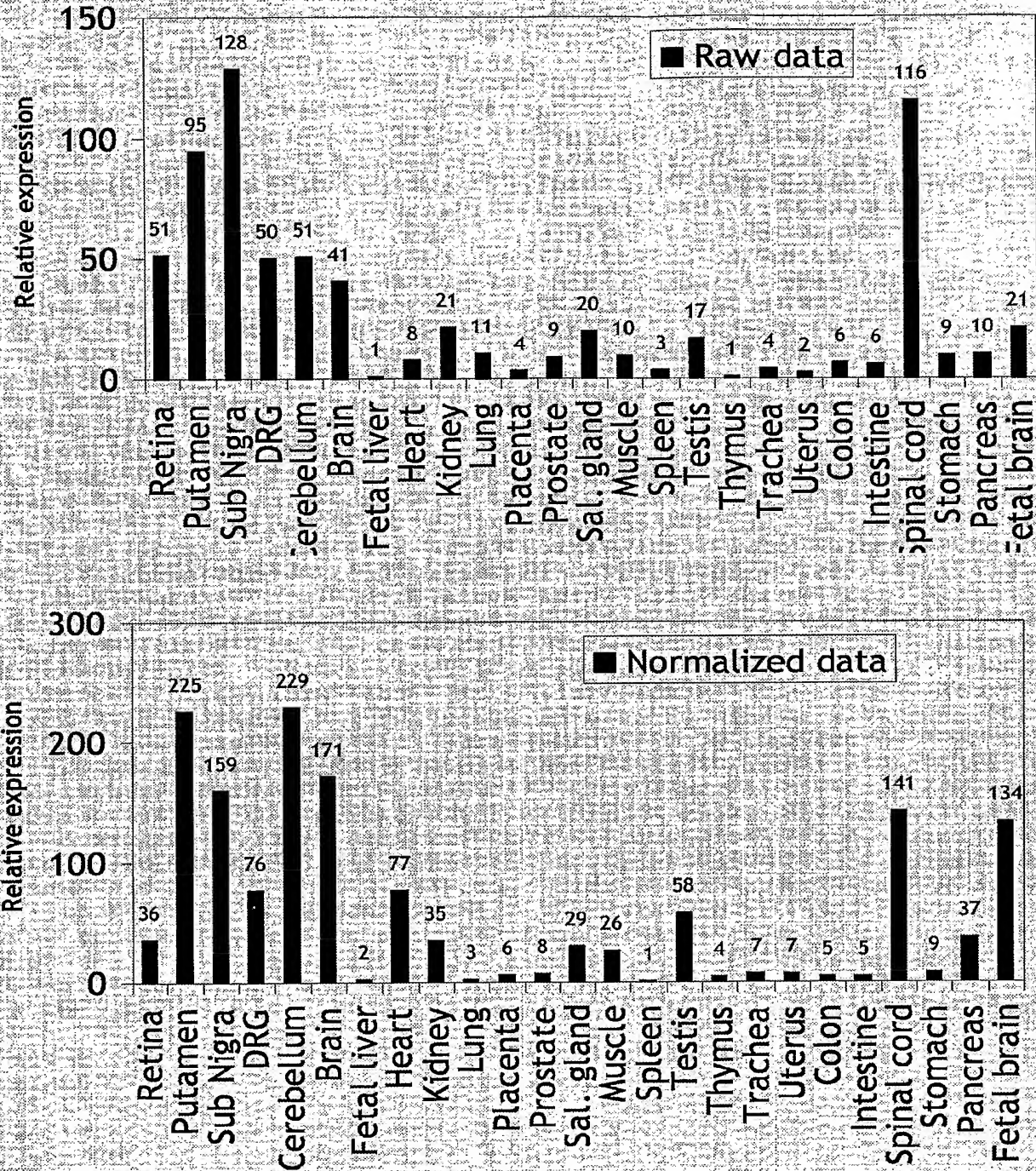
Mouse      LLAactsdfVIHGtiHGVahdTELQESVITVVVARVIRQTLPLFKESGSEGGQGRASIRTL 238
Rat        LTACTSDFVIHGtiHGVVHDMELQESVITVVATRIVIRQTLPLFQESGSEGRGQASVRTLL 238
Human      LLAactsdfVIHGtiHGVTHDVELQESVITVVAARVLRQTPPLFQAGRSGDQGLTSIRTPL 240
            *:***** *****.* *****.:*:*** ***: * * .:* :*:**

Mouse      RCGVRPGPGSFLFMGWSRFGEAWLGCAPRFQEFSSRVYSAALTTHLNPCEMALD 291
Rat        RCGVRPGPGSFLFMGWSRFGEAWLGCAPRFQEFSSRVYSAALAAHLNPCEVALD 291
Human      RCGVHPGPGTFLFMGWSRFGEARLGCAPRFQEFRRAYEAAARAHLHPCEVALH 293
            *****:***** ***** *.*.*:*.***:*.

```

4/13

Fig. 4



5/13

scoring matrix: BLOSUM50, gap penalties: -12/-2
 42.3% identity; Global alignment score: 747

```

      10      20      30      40      50 *
Innog. MRGAARAAWGRAGQPWPRPPAPGPPPPPLLLLLLLLAGLLGGAG-AQYSSDRCSWKGSGL
      :      ::      :      ::      :      :      :      :      :      :
NsG33  -----MGFPAAALLCALCCGLLAPAARAGYSEERCSWRGSGL
                        10      20      30

      60      70      *      80      90      100      *110
Innog. THEAHRKEVEQVYLRC AAGAVEWMYPTGALIVNLR-PNTFSPARHLTV CIRSFTDSSGAN
      :      :      :      :      :      :      :      :      :      :
NsG33  TQEPGS--VGQLALACAEGAVEWLYPAGALRLTLGGPDPR--ARPGIACLRPVRPFAGAQA
      40      50      60      70      80      90

      120     130     140      *      150     160     170
Innog. IYLEKTG-ELRLVDPDGDGRPGRVQC--FG-LEQGGLFVEATPQQDIGRRTTG FQYELVR
      ..      :      :      :      :      :      :      :      :      :
NsG33  VFAERAGGALELLLAEGPG-PAGGRCVRWGPERRALFLQATPHQDISRRVA AFRFELRE
      100     110     120     130     140     150

      180     190     200     210     220
Innog. RHRAS---DLHEL SAP--CRPCSDTEVLLAVCTSDFAVRGSIQQVTHEPERQDS AIHLRV
      :      :      :      :      :      :      :      :      :      :
NsG33  DGRPELPPQAHLGLGVDGACRP CSDAELL LAACTSDFVIHGIIHGVT HDVELQESVITVVA
      160     170*      *      180      *      190     200     210

      230     240     250     260     270     280
Innog. SRLYRQKSRVFEPVPEGDGHWQG--RVRTLLECGVRPGHGDFLFTGHMHFGEARLG CAPR
      ...      :      :      :      :      :      :      :      :      :
NsG33  ARVLRQTPLPLFQAGRSGD---QGLTSIRTPLRCGVHPGPGTFLFMGWSRFGEARLG CAPR
      220     230     240      *      250     260      *

      290     300      *      310
Innog. FKDFQRM YRDAQERGLNPCEVGT D
      :      :      :      :      :
NsG33  FQEFRRAYEAARA AHLHPCEVALH
      270     280     290

```

Fig 5

6/13

```

1  gcttgcgcg ggcgcggcg cgcgcgcgc cgcgcgcgc cgcgcgcgc cgcgcgcgc gacgccgcg
101  acgcgcgcgc cgcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc
    >>.....CDS.....>
    m g f p a a a l l c a l c c g l l a p a a r a g y s e
201  ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc
    >.....CDS.....>
    e r c s w r g s g l t q e p g s v g q l a l a c a e g a v e w l y p
301  ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc
    >.....CDS.....>
    a g a l r l t l g g p d p r a r p g i a c l r p v r p f a g a q v
401  ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc
    >.....CDS.....>
    f a e r a g g a l e l l l a e g p g p a g g r c v r w g p r e r r
501  ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc
    >.....CDS.....>
    a l f l q a t p h q d i s r r v a a f r f e l r e d g r p e l p p q
601  ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc
    >.....CDS.....>
    a h g l g v d g a c r p c s d a e l l l a a c t s d f v i h g i l
701  ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc
    >.....CDS.....>
    h g v t h d v e l q e s v i t v v a a r v l r q t p l f q a g r
801  ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc
    >.....CDS.....>
    s g d q g l t s i r t p l r c g v h p g p g t f l f m g w s r f g e
901  ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc
    >.....CDS.....>
    a r l g c a p r f q e f r r a y e a a r a a h l h p c e v a l h -
1001 ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc
1101 aaaaaaaaaa

```

Fig 6

7/13

```

1  ccacgcgtcc gccacgcgt ccgcgccttct ttgcgcgctc tgttggggcc tcttggccgc gtccgctcac gctgggtact cggaagaccg ctgcagctgg
>>.....CDS.....
101  h a s a h a s a l l c a l c c g l l a a s a h a g y s e d r c s w
aggggcagcg gttgaccca ggagcctggc agcgtggggc agctgaccct ggaactgtact gagggcgcta tcgagtggct gtaccagct ggggcgctgc
>.....CDS.....
201  r g s g l t q e p g s v g q l t l d c t e g a i e w l y p a g a l
gactgaccct gggcgccccc gatccgggca cacggcccag catcgtctgt ctggcccag agcgccctt cgctgggtgc caggtcttgc ctgaacgtat
>.....CDS.....
301  x l t l g g p d p g t r p s i v c l r p e r p f a g a q v f a e r
gacgggcaat ctagagtgc tactggccga gggcccggac ctggctgggg gcgctgcac ggcgtgggt cccgcgagc gccgagccct ttctctgcag
>.....CDS.....
401  m t g n l e l l l a e g p d l a g g r c m r w g p r e r a l f l q
gccacaccac accgcgacat cagccgcaga gttgctgcct tccgttttga actgcacgag gaccaacgtg cagaaatgc tccccaggct caaggtcttg
>.....CDS.....
501  a t p h r d i s r r v a a f r f e l h e d q r a e m s p q a q g l
gtgtggatgg tgctgcagg ccctgcagt atgccgagct cctcctggct gcatgcacca gtgattttgt gatccacgg accatccatg gggtcgcca
>.....CDS.....
601  g v d g a c r p c s d a e l l l a a c t s d f v i h g t i h g v a
tgacacagag ctgcaagaat cagtcacac tgtgtgtgtt gctcgtgtca tcggccagac actgccactg ttcaagggaag ggagctcga gggccaaggc
>.....CDS.....
701  h d t e l q e s v i t v v a r v i r q t l p l f k e g s s e g q g
cgggcctcca ttctgacctt gctgcgctgt ggtgtgcgtc ctggcccagg ctctctctc ttcattgggt ggagccgatt tggcgaagct tggctgggt
>.....CDS.....
801  x a s i r t l l r c g v r p g p g s f l f m g w s r f g e a w l g
gtgtccccg ctccaagag ttacgccgtg tctattcagc tgctctcag accatctca accoatgtga gatggcactg gactgagaga cctgggagca
>.....CDS.....
901  c a p r f q e f s r v y s a a l t t h l n p c e m a l d -
agccctggat ggaccttctt ctggagatgg ggtgttggg aggtgatgg gaggtgggtt tggctcggat ggcatoctgg taccacagat
1001 gagctggtag aatactaagt aatctggacc ataaaaaaaa aaaaaaa

```

Fig 7a


```

1  gggcagccgc gccgcgggct gctcgcgctg cggccccgac cctccggggg cagcagtcog agggcccgcc gcgtccoccta accatgctgg
   >>CDS.>
   m l
91  tagccacgct tttttgcgog ctctgttgog gcctcctggc cgggtcogct caogctggct actoggaaga ccgctgcagc tggaggggca
   >.....CDS.....>
   v a t i l c a i l c g i l a s a h a g y s e d r c s w r g
181  ggggtttgac ccaggagcct ggcagcgtgg ggcagctgac cctggactgt actgagggcg ctatcgagtg gctgtacca gctgggggog
   >.....CDS.....>
   s g l t q e p g s v g q i t l i d c t e g a i e w l y p a g a
271  tggcctgac cotggggcgc cccgatccgg gcacacggcc cagcatgctc tgtctgcgc cagagcggcc ctctcgctgg gccaggtct
   >.....CDS.....>
   i r i t l g g p d p g t r p s i v c l r p e r p f a g a q v
361  tggctgaacg tatgaccgac aatctagagt tgctactggc cgaggggccg gacctggctg gggggccgctg catgcgctg ggtcccccg
   >.....CDS.....>
   f a e r m t g n l e i l l a e g p d l i a g g r c m r w g p r
451  agggccgagc ccttttctg caggccacac cacacggcga catcagcgc agagttgctg ccttcgctt tgaactgcac gaggaccaac
   >.....CDS.....>
   e r x a l f l q a t p h r d i s r r v a a f r f e l h e d q
541  gtgcagaaat gtctcccccag gctcaagtc ttggttgga ttgtgctgc aggcctgca gtgatgcga gctcctctg gctgcagtca
   >.....CDS.....>
   r a e m s p q a q g l g v d g a c r p c s d a e l l i a a c
631  ccagtgattt ttgtatccac gggaccatcc atggggtgc ccatgacaca gagctgcaag aatcagtcac cactgtggtg ttgctcgtg
   >.....CDS.....>
   t s d f v i h g t i h g v a h d t e l q e s v i t v v a r
721  tcacgcgcca gacactgcca ctgttcaagg aaggagctc ggaggggcaa ggccgggct ccattcgac ctgctgcgc ttggtgtgctg
   >.....CDS.....>
   v i r q t l p l f k e g s s e g q g r a s i r t l l r c g v
811  gtcctggccc aggtccttc ctcttcacg gctggagcog atttgogaa gcttggtgctg gctgtgctcc ccgcttccaa gatttcagcc
   >.....CDS.....>
   x p g p g s f l f m g w s r f g e a w l g c a p r f q e f s
901  gtgtctattc agctgtctc acgaccatc tcaaccatg tgagatgga ctggactgag agacctggga gcaagccctg gatggacctt
   >.....CDS.....>
   r v y s a a l t t h i n p c e m a l d -
991  ctcttggaaga tgggtgtgtg gggaggggtga tgggaggtg ggtgagaagg gtgtggtctg gatggcatcc tgggtaccac agtgagctg
1081  tagaatacta agtaatotg accataccag ccactgtagt catggtcttc ttgtggcaggc agcatacca gctctgtgcc tgcctcactt
1171  tgtctactct ccagctgtct gcccttctaa cccttcttag cctgtgtaac agtgagctca tgttttctc gaattccagg gtgctgctg
1261  ggttcagagc aacgctgccc tagtttgga gacttgagct aattgtttt tttttgtttg tttttttgtt tgttttaagg tggcctgggg
1351  gggggggcaa aca

```

Fig 7b

9/13

```

1  atgctggttag cggcgcttct ctgcgcgctg tgctcgggc tcttggtgc gtccgctcga gctgggtact ccgaggaccg ctgcagctgg aggggcagcg
>>.....CDS.....>
    m l v a a l l c a l c c g l l a a s a r a g y s e d r c s w r g s
101 gtttgacca ggaacctggc agctggggc agctgacct ggattgtact gaggtgcta tccagtggct gstatccagct gggcgctgc gcctgactct
>.....CDS.....>
    g l t q e p g s v g q l t l d c t e g a i e w l y p a g a l r l t
201 aggggctct gatccgggca cgcggcccag catcgtctgt ctgcgccaa cagggccctt ccgtgggtgc caggtcttcg ctgaacggat ggcgggcaac
>.....CDS.....>
    l g g s d p g t r p s i v c l r p t r p f a g a q v f a e r m a g n
301 ctgagttgc tactggcca gggccaaggc ctggctggg gccgtgcat gcgctgggt cctgcgagc gccgagccct ttctctgcag gccacggcac
>.....CDS.....>
    l e l l l a e g q g l a g g r c m r w g p r e r r a l f l q a t p
401 accgggacat cagccgcaga gttgctgct tocaatttga actgcacgag gaccaacgtg cagaaatgc tccccaggcc caaggttttg gtgtggatgg
>.....CDS.....>
    h r d i s r r v a a f q f e l h e d q r a e m s p q a q g f g v d
501 tgcctgcagg cctgcagtg atgcgagct ccttctgact gcatgcacca gtgactttgt gatccatggg accatccatg gggctcgtcca tgacatggag
>.....CDS.....>
    g a c r p c s d a e l l l t a c t s d f v i h g t i h g v v h d m e
601 ctgcaagaat cagtcacac tgtgttgcc actcgtgtca tccgccagac actgccactg ttccaggaag ggagctcga gggccggggc caggcctccg
>.....CDS.....>
    l q e s v i t v v a t r v i r q t l p l f q e g s s e g r g q a s
701 ttctgacctt gttgcgctgt ggtgtgcgtc ctggcccagg ctcttctc tccttctc ttcatgggt ggagccgatt tggcgaagct tggctgggt gcgctcccg
>.....CDS.....>
    v r t l l r c g v r p g p g s f l f m g w s r f g e a w l g c a p
801 ctccaagag ttacgacctg tctattcagc tgctctcgc gccacactca acccatgtga ggtggcactg gactgagaga cctgggagca agccctggat
>.....CDS.....>
    r f q e f s r v y s a a l a a h l n p c e v a l d -
901 ggatcttct ctggggatgg ggtgttggg aggggtgata ggagggtggg tgggaagggt gtggtcaga tggtatcctg gtaccacag tgagggtggt
1001 gaatactaaa taacctggat cacacc

```

Fig 8

10/13

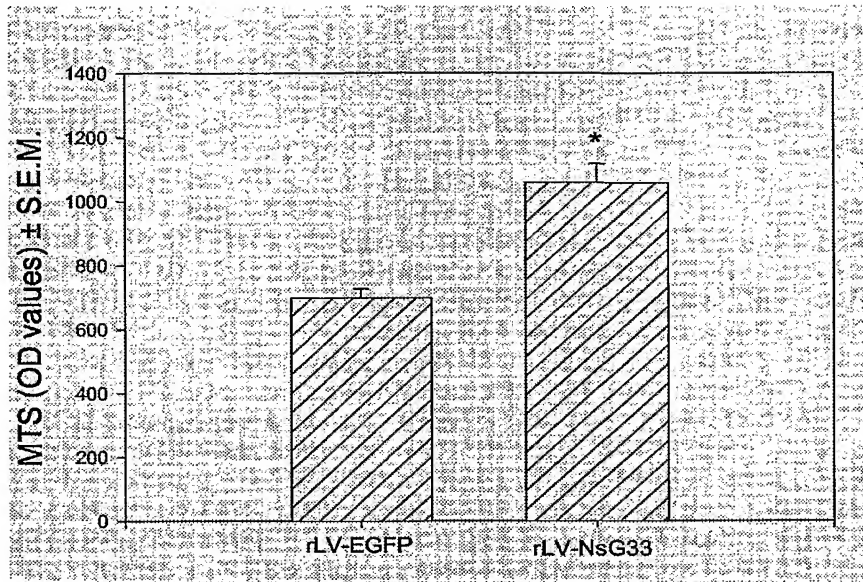


Fig. 9

11/13

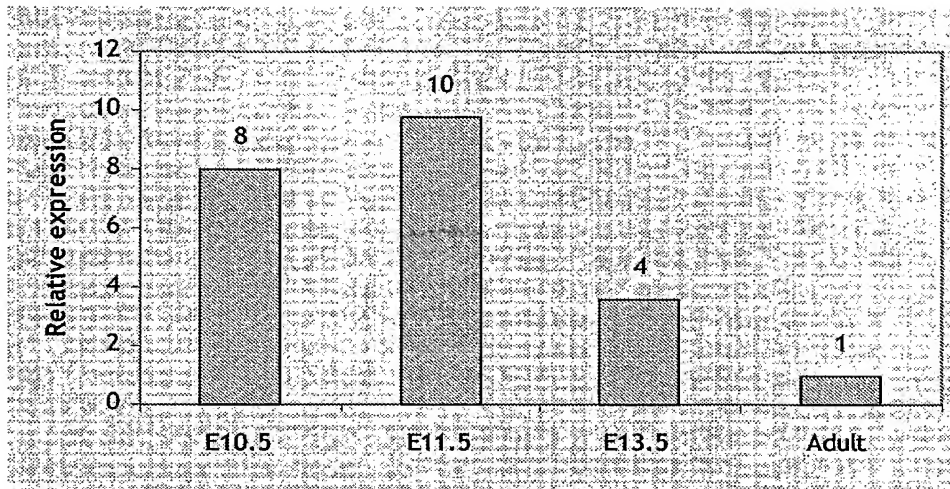


Fig. 10A

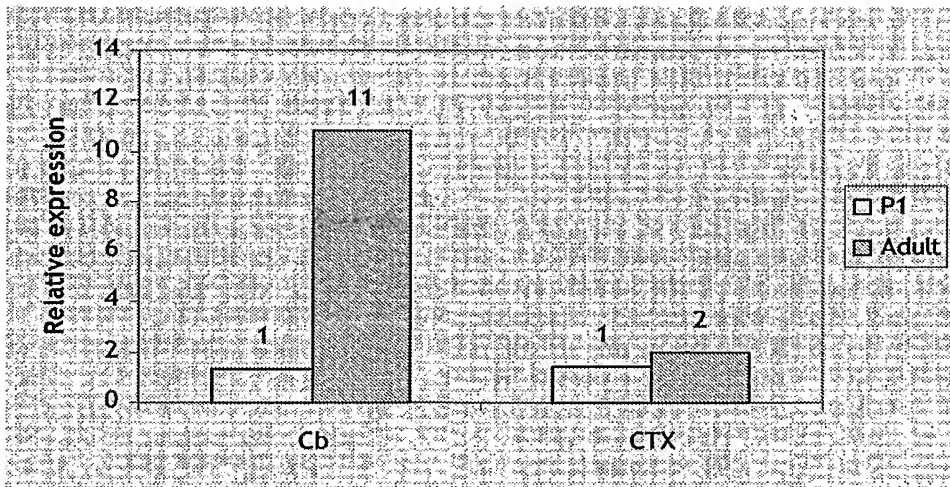


Fig. 10B

12/13

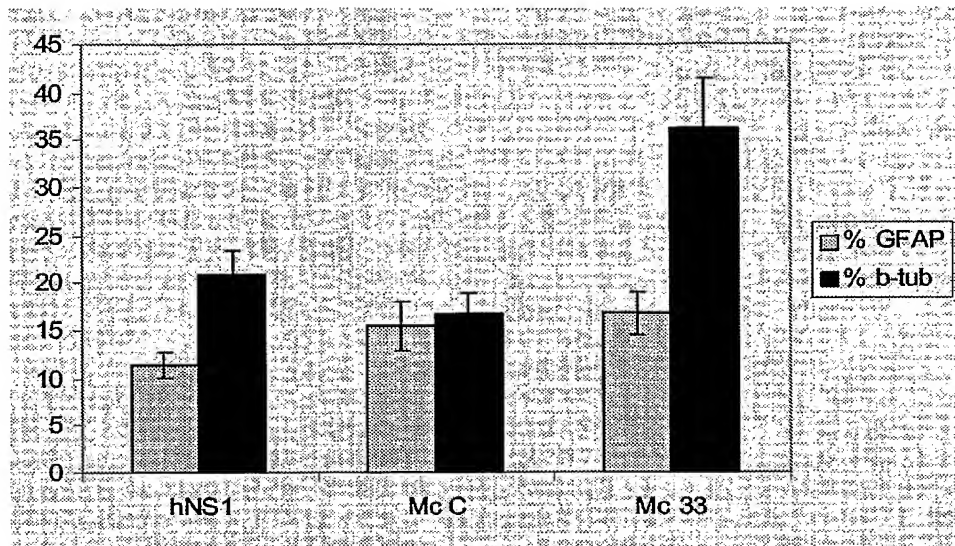


Fig. 11A

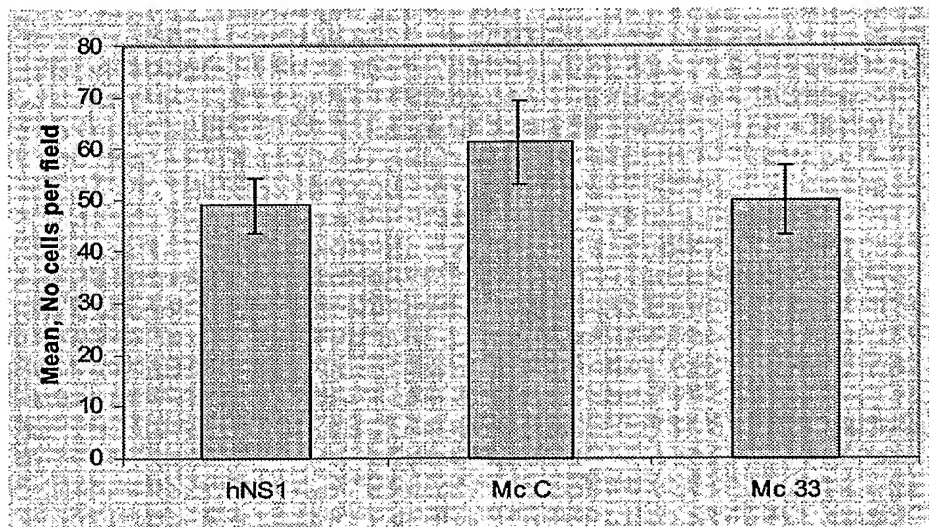


Fig. 11B

13/13

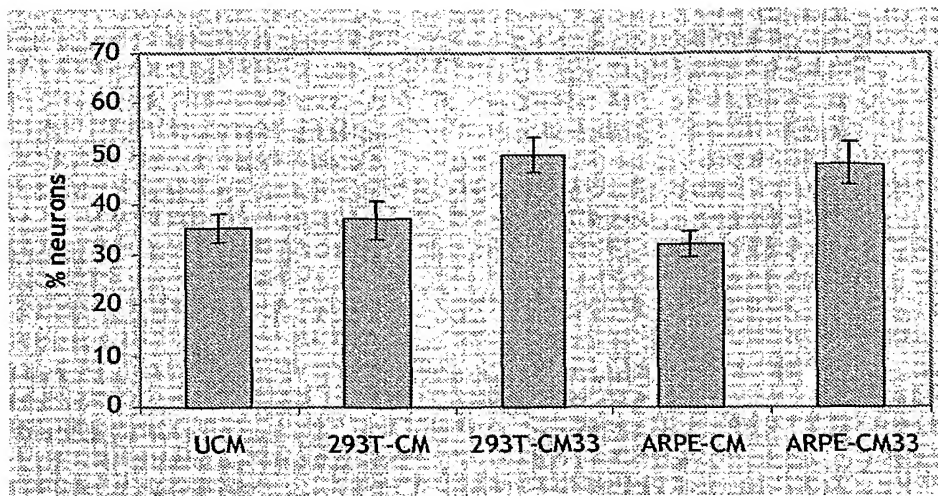


Fig. 12